Txtng in three European languages: does the linguistic typology differ?

Rachel Panckhurst

Praxiling UMR 5267 CNRS,
Université Paul-Valéry Montpellier 3
Route de Mende
34199 Montpellier cedex 5
France

rachel.panckhurst@univ-montp3.fr

[4661 words total ; abstract 120 words]

Txtng in three European languages: does the linguistic typology differ?[1]

Rachel Panckhurst,
Praxiling UMR 5267 CNRS, Université Paul-Valéry Montpellier 3
rachel.panckhurst@univ-montp3.fr

Abstract: We have conducted research on computer-mediated communication within a French University since 1996. Discourse appearing in email messages, forums, chat sessions is shaped in a particular way, precisely because one uses a computer. The computer becomes a tool, a sort of mediator, indirectly modifying the discourse. A new discourse 'genre' which we call *mediated electronic discourse* is created. Our research has recently included texting or SMS-type writing strategies either through computer or mobile telephone usage. Based on French and Belgian research, we have devised a new typology for French SMS writing (Panckhurst, 2009). Student corpora comparing French, Spanish and Italian allows plurilingual verification in order to reveal any variations. SMS and "dialogical interactions" are the next step for this research.

Keywords: SMS, txtng, computer-mediated discourse, mediated electronic discourse.

1. Introduction

1.1. Mediated electronic discourse

Mediated electronic discourse (MED) (including electronic mail, forums, chats, blogs, etc.) is characterised by several phenomena which appear below (*cf.* Panckhurst 2006a, b, 2007, Véronis & Guimier de Neef, in Sabah, 2006).

Main MED features:

- *smileys* to introduce non-verbal semiological aspects; specific typography: words in uppercase, lengthening or repetition of letters, (which, in certain cases may simulate intonation, and

[1] We would like to thank the following students, friends & colleagues who have provided useful information on French, Spanish and Italian texting usage: Dorian Boudrique, Hélène Catapano, Mark Chu, Flavia Coassin, Stéphanie Doston, George Ferzoco, Delphine Fréjaville, Sonia Pollet, Teresa Stabile, Elisabetta Tortorella, Bertrand Verine.

therefore indicate some paraverbal information), marks such as '>'or '|' (indicating a repetition of discourse between sender and recipient);

- spelling, grammatical mistakes and absence (or reduction) of punctuation (*cf.* Panckhurst, 1998; Véronis & Guimier de Neef, in Sabah, 2006);
- neology or neography (*cf.* Véronis & Guimier de Neef, in Sabah, 2006), for instance, SMS abbreviations or words borrowed from foreign languages.

More linguistic features include:

- predominant usage of the present tense (often over 60-70%) as opposed to imperfect/past, future, conditional, imperative;
- high usage of first person deictic pronouns (as compared to second and third person pronouns);
- lower percentage of verbs (often under 20-25%) compared to other more traditional written forms (over 20-25%), and among verbs used, frequent usage of modals (between 20 and 30% of overall verb usage);
- increased usage of ellipsis (for instance: *Vous remerciant/Thanking you*; *Impossible de trouver le document à enregistrer sur clef USB*/*Impossible to find the document to save on flash drive*)

Other more extra-linguistic aspects which are typical of online communication include:

- relational: conciseness, rapidity, anguish/worry (if a long silence is observed before responding to messages), aggressiveness, impulsiveness, an (illusionary) impression of proximity, protective barriers (no direct face-to-face contact), etc.
- communication context: reduction or absence of introductions and closures, non-observance of conversational rules (turn-taking, floor-taking, adjacency pairing, etc).

More recently (2005-2008), we have expanded our research to include mobile-telephone SMS messages (*cf.* Panckhurst, 2009). The following table shows how syntactic categories vary greatly according to the type of communication tool used:

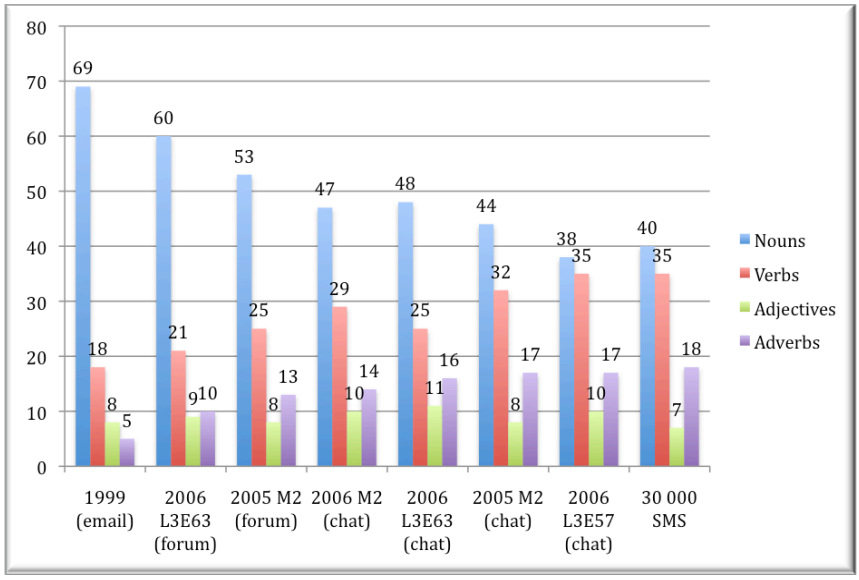

Table 1: Percentage of syntactic categories (tokens) used between 1999 and 2006 (email, forum, chat, SMS)

Until 2004, our corpora showed that syntactic category usage for nouns, verbs, adjectives and adverbs used in MED situations was similar to more conventional written forms (i.e. a high percentage of nouns and few verbs). However, from 2005 up until now, a radical change has occurred: noun usage has decreased, adjectives remain stable, but verb and adverb usage has increased dramatically. This shows a significant linguistic evolution: MED can be either more closely linked to written forms, or oral forms, according to the communication tool used. In Table 1, chats and SMS usage[2] appear to be more "oral" whereas email and forums remain linked to more "written" forms.

1.2. eSMS or txtng

Following on from the above classification for MED, analysis of SMS or txtng-style writing (in French, I have created the term *eSMS* for "*écriture* de type SMS" or "écrit SMS" (*cf.* Panckhurst, 2009), which I shall continue to use here), offers new perspectives for the study of language in evolution:

> "Some people dislike texting. Some are bemused by it. Some love it. I am fascinated by it, for it is the latest manifestation of the human ability to be linguistically creative and to adapt

[2] The right-hand column in Table 1 shows results from analysis of SMS data. We submitted 30,000 text messages which had been transcribed into standard French (see Fairon *et al*, 2004-2007) to a French computational linguistics tool (Cordial, Synapse) and we then extracted results for noun, verb, adjective and adverb usage.

> language to suit the demands of diverse settings. In texting we are seeing, in a small way, language in evolution". Crystal (2008, p. 175)

Neology (lexical creativity, essentially foreign languages and colloquial/slang usage (also "verlan" in French) and *neography* (innovative spelling) are of the upmost importance as they are features which are frequently used in *eSMS*.

In this article, a typology for French is presented (Panckhurst, 2009) and is then confronted with data from Italian and Spanish. This allows plurilingual verification in order to reveal any variations, and subsequently adjust the typology accordingly.

2. *Typology*

In previous research (Panckhurst, 2009), we have justified devising a new typology for French eSMS following on from other proposals (Anis, 2004, Fairon *et al.* 2006b, Liénard, 2007, Véronis et Guimier de Neef, 2006) for several reasons. These are mainly linked to: varying terminology for identical phenomena, complex category confusions, (phonetic, graphical, etc.) descriptions which are placed on the same level as the resulting meaning/interpretation of the messages, sole categories which are used to highlight a particular phenomenon but include examples which can also belong to another category, etc. As these aspects are detailed in the previous research, we prefer directly presenting our typology for French here and then confronting it with two other languages: Spanish and Italian.

In the following typology for French we exclude *neology*. *Neography* is first presented feature by feature (see Table 2), and then combinations are included (*cf.* § 2.1). This is in order to avoid confusion as eSMS often includes complex features.

2.1. French eSMS typology

I. Substitutions
1. phonetic: **1. entire**: a sound is replaced by one character (letter or number) The spelling of the lexeme is entirely modified: *o* (**eau**), *7* (**cet**). **2. partial**: replacement of digrams and trigrams, corresponding to phonemes. Spelling of the lexeme is partially modified: *ossi* (**au**ssi), *allé* (all**er**), *bo* (b**eau**) ; intervocalic "s": *bizes* (bi**s**es) 3. with variation: *bisoo* (biso**u**)
2. graphical: **1. elision, typography, capitals/lower case**: replacement of the French apostrophe or hyphen by spaces: « m en » (m'en), « est ce que » (est-ce que) ; capitalising whole messages or substituting capitals with lower case or vice versa. **2. icons, mathematical symbols, special characters, rebuses, pictograms, logograms**: (*, + = => @); *à* + (à **plus**), *de grandes* @ (de grandes **oreilles**) **3. with variation :** *bisoux* (bisous) *mwa* (moi)
II. Reductions
1. phonetic: **1. morpho-lexical shortenings:** **1.1. truncations** : *ordi* (ordinateur, apocope), *'lut, Net* (salut, Internet, aphaeresis). **1.2. initialisms (alphabetisms & acronyms): all letters are omitted except the first**: *ASV* (âge, sexe, ville), *mdr* (mort de rire), *tvb* (tout va bien), *tlm* (tout le monde), *lol* (laughing out loud). **2. variation** : ui (*oui*), i (*il*)
2. graphical: **1. suppression of mute 1) word-endings 2) beginnings**: 1) *échange* (échange**s**), *vou* (vou**s**), *peu* (peu**t**) , *chian* (chian**t**), *fou* (« m'en fou**s** »), **drop of the unstable *e***: *douch* (douch**e**); 2) *ôtel* (**h**ôtel) **2. consonant contractions/clippings & abbreviations**: *dc* (donc), *pr* (pour), *ds* (dans); **double consonants**: *ele* (elle), *poura* (pourra) ; **semantic abbreviations/initialisms** (abbreviations reduced to initials): *t* (te/tu) *p* (peux/pas). **3. agglutinations :** *jattends* (j'attends).
III. Suppression/absence or rarefying:
1. graphical: **1. typography & punctuation** **2. diacritic signs** : *ca* (ça), *voila* (voilà)
IV. Augmentations & additions :
1. graphical: **1. repetition of characters and/or punctuation** : *suuuupppeeerrr !!!!!!* **2. semiological representations (Smileys/emoticons)** :-) **3. character increase**: *oki* (ok), *les zamours* (les amours) **4. onomatopoeia**: *mouarf, arfff, bof.*

Table 2: typology of simple eSMS phenomena for French.

2.1.1. Complex phenomena. Table 2 is insufficient for dealing with all cases of neography, as, in many instances, eSMS involves combinations of two or more graphical and/or phonetic substitutions, reductions, suppressions, rarefying and additions. Some examples follow:

- agglutination (II.2.3) + suppression of mute word-endings (II.2.1) + entire phonetic substitution (I.1.1) : *7éta* "cet état";
- multiple entire phonetic substitution (II.2.1) : *2manD* (demander), *6T* (citer), *A2M1* (à demain);
- character increase (IV.1.3.) and graphical substitution with variation (I.2.3.): *mwouah, moua* (moi)

A recent student study (Fréjaville & Doston, 2009), conducted on analysis of 300 text messages exchanged between people aged between 17 and 32 in French, has shown that the 4 most frequent combinations (including 2 or 3 phenomena), in descending order, are as follows:

(1) substitution & reduction:

substitution (phonetic, partial) & reduction (graphical, agglutination): *javé* (avais > avé; j' > j); *kon* (qu'on: qu' > qu > k);

substitution (phonetic, partiel) & reduction (graphical, mute word-endings): *vrémen* (vraiment: ai > é; t > ø);

(2) reduction & reduction:

reduction (graphical, initialism) & reduction (graphical, agglutination): *sfera* (se > s; se fera > sfera); *jpense* (je > j; je pense > jpense);

(3) substitution & reduction & reduction:

cke (c'est que: que > ke (phonetic, partial); c'est > c (graphical, semantic abbreviation); c'est que > cke (graphical, agglutination);

jfiniré (je finirai: ai > é (phonetic, partial); je > j (graphical, initialism); je finirai > jfiniré (agglutination);

ti (t'y: y > i (graphical, with variation); te > t (graphical, semantic abbreviation); t'y > ti (agglutination)

(4) reduction & reduction & reduction:

reduction (graphical, semantic abbreviation) & reduction (graphical, mute word-endings) & reduction (agglutination): *ta* (tu as: tu > t; as > a; tu as > ta; *jdevien* (je deviens: je > j; deviens > devien; je deviens > jdevien);

2.2 French eSMS compared to Spanish and Italian

In order to check the validity of the typology for French eSMS in comparison to Spanish and Italian, two other student studies were initiated (Boudrique *et al,* 2008 — analysis of 84 Spanish text messages; Stabile & Tortorella, 2008 — analysis of 208 Italian text messages). The following results (see Table 3, parts 1-3) show that the features stipulated in Table 2 can generally also be used in these languages. However, several specific situations do differ.

2.2.1. Substitution of trigrams. As pronunciation and spelling are much closer in Spanish and Italian compared to French, it is not surprising that trigram simplification does not appear in those languages, as it does in French: "beau" > "bo".

2.2.2. Phonetic substitution: double to single letter modification: bu > w; ll > y; ch > x (Spanish). Phonetic variation seems more popular in the Spanish and Italian corpora than in French. Some phenomena may be of interest. For instance, the letter "w" is very rare in standard Spanish, only normally existing in foreign words. It is however also used in some regions/countries and in varying sociolects. In many eSMS examples, "bu" is replaced by "w": *bueno > weno; guapa > wapa,* this corresponds to an economical measure when using the mobile telephone. Even if the letter "w" is rarely used in standard Spanish, the corresponding pronunciation does exists: *huera* is normally pronounced *wera*. This tendancy needs to be confirmed with intercultural differences; for instance, in other (Latin American) countries, a word like *huera* may more often be pronounced *gwera*.

The double "ll" at the onset of a word is fairly systematically replaced with "y" since the pronunciation is fairly similar (in Argentina, for instance, "ll" is systematically pronounced "y") and the economical character of the mobile telephone keyboard is a factor (*llave > yave, llamo >yamo*). It is not necessarily the case for intervocalics; the following words were written as such in the Spanish corpus: "ella", "allí", "pollito". "Ch" >"x" is borrowed from catalan, in which "x" is pronounced "ch" (*chiquitina > xiquitina*). This enables the user to press one instead of two keys on the mobile phone.

Single letter modification (Spanish). Although "v" is often pronounced "b" in standard Spanish (there are of course geographical variations for this), in eSMS the replacement is not often used, probably because the key-pressing on the mobile phone is almost the same. The following example was the only one found in the whole Spanish corpus:

*Pdrs **b**nir a brme ?* (¿Podrás venir a verme ?; Will you be able to come and see me?)

2.2.3. Suppression of final characters. Surprisingly enough, single characters of (pronounced) word-endings are sometimes suppressed; however, in the Italian corpus this only seems to occur in dialectal situations (Napolitano in the present instance), where the full word is not pronounced in oral Italian in the first place; therefore this situation may be identical in spoken dialectal Italian and eSMS: *scem* (scema), *avert* (averti) (Italian). This may be similar in some regions of Spain (Andalucia) where pronounced word endings are omitted orally (final "s" dropped: *eres >ere, vamos >vamo,* etc.). Regional corpora should allow verification of these phenomena.

2.2.4. Mute word-endings (French)/beginnings (French, Spanish, Italian). In the section on reductions, "suppression of mute word-endings" is non applicable for Spanish and Italian, as this phenomenon does not exist in either language, simply because all characters (apart from the "h" for Spanish/Italian) are pronounced.

In Spanish, mute word-beginnings may be suppressed in the situation of the silent "h" (one must note, however, that in certain regions and countries "s" is pronounced as an aspirate "h"). For instance : *hermanos* (brothers) is often written either "ermanos" or "rmanos" (even if the latter possibility should technically be proncounced "erre"+"manos") in eSMS; "hemos" is also written "emos" in eSMS. These are the only two examples encountered in the Spanish corpus. In the Italian corpus, however, all instances of the mute "h" appear intact, as such (*ha, ho, hai, hanno*).

2.2.5. Apostrophes & agglutination (French). Apostrophes are not used in Spanish, so for instance, there is no elision possible for "me acuerdo" and it would therefore not be shortened to "*m'acuerdo". In French, of course, it is very frequent: "m'aime", "j'aime" but also with determiners and nouns: "l'apostrophe". In Italian, apostrophes are used with reflexive verbs, clitic pronouns or determiners + nouns, etc. The agglutination factor often revealed in French eSMS seems less frequent in Spanish or in Italian: *jattends* (I'm waiting), *jy* (j'y, as in j'y vais: I'm going there), *nai* (n'ai, as in je n'ai pas: I don't have). In the Italian corpus, not one example of apostrophe reduction combined with agglutination is apparent, but this may need to be confronted with further corpora.

In French, agglutination in eSMS also appears in situations where apostrophes are not used in standard French, for instance: *jregarde* (je regarde), *jtravaille* (je travaille), etc. In Spanish, no agglutination examples using pronominals or reflexives (subject pronouns are of course not used in unstressed Spanish or Italian) appear in the corpus; a space is systematically inserted with initialism abbreviations: *m aburro* (me aburro), *s nvian* (se envian). In fact, very few examples of agglutination appear in the Spanish corpus; the three following were observed: 1) prepositions & nouns (*dtus padres* > de tus

padres); 2) interrogative adverb & verb combinations (*d donderes ?* > ¿De dónde eres ?); 3) expressions with or without subject pronouns: *noseke* (no sé qué), *yokese* (yo qué sé). In the Italian corpus, the only example of agglutination is: *vabbè* (va bene), but this also involves consonant doubling and apocope, and since it is typical in oral Italian, it is not a particular feature of eSMS.

2.2.6. Double consonant simplification (Spanish & Italian). In both Spanish and Italian, usage of double consonants is less frequent than in French. In French the simplification is frequent: *elle >ele* (she); *pourra >poura* (will be able to) In Spanish, the main consonants which are doubled are: *ll, rr, cc*, and *nn*. Since the "ll" is often replaced by "y" (word-beginnings), that leaves three other main possibilities: the only double "nn" encountered in the corpus is when *mañana* (tomorrow) is condensed to *mñn*, which is not once simplified to *mñ*; 1 occurrence of the double "rr" is encountered: *arreglo*, but without simplification; the double "cc" is not apparent.

2.2.7. Diacritic signs and punctuation (Spanish). Diacritic signs are often used in written Spanish and these are often suppressed in eSMS. This may lead to ambiguity, as certain words can be indirect exclamations or interrogatives, depending on the context: *quien/quién* (who), *donde/dónde* (where), *cuando/cuándo* (when). Initial punctuation, which does not exist in either French or Italian, is systematically suppressed (¡, ¿) in Spanish text messages.

2.2.8 Complex phenomena variations (Spanish & Italian)

In both Spanish and Italian, consonant doubling is used in combination with initialisms: *gg* (giorno, Italian), *aa* (años, Spanish), *dd* (días, Spanish). This does not seem to have an equivalent in French, where double consonants usually appear in consonant contractions, for instance: *qq* which may sometimes be used for "quelques", *bb* for "bébé", *cc* for "coucou", *mm* for "meme", *nn* for "non", *pp* for "pépé", *ss* for "sans", *tt* for "tout", etc. However both consonants exist in the original words, which differentiates the case from both Spanish and Italian.

In Spanish, substitution may originate from mainstream Spanish (*todos* (all) becomes *to2* — since "dos" also means "two") or in some cases is possibly influenced by Catalan (*adios* (goodbye) becomes *a2* (because in Catalan "déu" means God and is similar to the plural "dues" meaning "two" — for a foreigner this may be confusing, as *a2* could be interpreted as meaning "ados" which of course doesn't exist).

2.3. Trilingual eSMS typology: French, Spanish, Italian

I. substitutions

1. phonetic

1. entire: a sound is replaced by one character (letter or number); the spelling of the lexeme is entirely modified

French: *o* (**eau**), *7* (**cet**); possible in complex neology: *6T* (citer)
Spanish: *c* (**sé**), *s* (**es**); possible in complex neology: *to2* (to**dos**), *a2* (a*dios* from *adéu*, possible Catalan influence)
Italian: *c* (**ce**), 6 (**sei**=pronoun, *you*); possible in complex neology: *3no* (**tre**no)

2. partial: replacement of 1) digrams and 2) trigrams, corresponding to phonemes; spelling of the lexeme is partially modified

French: *1) ossi* (**au**ssi), *allé* (all**er**); *2) bo* (b**eau**); intervocalic "s": *bizes* (bises)
Spanish: 1) *yamar* (**ll**amar), *yave* (**ll**ave), *chikitina* (chi**qu**itina), *xiquitina* (**ch**iquitina, from catalan); 2) N.A.; intervocalic "s": *bezos* (be**s**os)

Italian: 1) *kiudere* (**ch**iudere); 2) N.A.

3. with variation:

French: *bisoo* (biso**u**)

Spanish: *weno* (**bu**eno), *wapa* (**gu**apa)

Italian: *poxo* (po**ss**o) — from Latin: "x" became "ss".

2. graphical

1. elision, typography, capitals/lower case: 1) replacement of apostrophe or hyphen by spaces; 2) capitalising whole messages or substituting capitals with lower case or vice versa

French: 1) *m en* (m**'**en), *est ce que* (est**-**ce que); 2) upper/lower case substitution apparent
Spanish: 1) no hyphen or apostrophe in literary Spanish; 2) upper/lower case substitution apparent
Italian: 1) no hyphen in Italian; l ho (l**'**ho)2) upper/lower case substitution apparent

2. icons, mathematical symbols, special characters, rebuses, pictograms, logograms: (*****, + = => @)

French: *à* + (à **plus**), *de grandes @* (de grandes **oreilles**); partial substitution not encountered in French corpora

Spanish: + (**más**), - (**menos**), @ (email), x (**por**); also partial substitution: *xq* (**por**que), *xo* (**per**o), *xa* **(par**a)
Italian: + (**più**), - (**meno**), x (**per**); also partial substitution: *xso* (**per**so), *xciò* (**per**ciò), *xché* (**per**ché)

3. with variation

French: *bisoux* (bisou**s**) *mwa* (m**oi**);

Spanish: not encountered in corpus

Italian: not encountered in corpus

Table 3 (part 1): typology of simple eSMS phenomena across three languages: French, Spanish, Italian

II. reductions

1. phonetic

1. morpho-lexical shortenings

1.1. truncations: 1) apocope; 2) aphaeresis

French: 1) *ordi* (ordinateur); *2) 'lut, Net* (salut, Internet)

Spanish: 1) *peli* (pelicula); 2) *stoy* (estoy), *stas* (estas)

Italian: 1) *scem* (scema), *avert* (averti), *veng* (vengo), *capi* (capito), *appunt* (appuntamento); 2) *na* (una), *sto* (questo)

1.2. initialisms: 1) alphabetisms & 2) acronyms: all letters are omitted except the first of each word

French*:* 1) *ASV* (âge, sexe, ville), *mdr* (mort de rire), *tvb (tout va bien), tlm* (tout le monde); 2) *lol* (laughing out loud) sic., used in French

Spanish: 1) *tqm* (te quiero mucho); 2) vab (¿ Vienes a buscarme ?)

Italian: 1) tvb (ti voglio bene);

2. with variation

French: *ui* (oui), *i* (il)

Spanish: only encountered in complex neology: *a2* (adios)

Italian: not encountered in corpus

2. graphical

1. suppression of mute 1) word-endings; 2) word-beginnings

French: *échange* (échanges), *vou* (vous), *peu* (peut), *chian* (chiant), *fou* (m'en fous); drop of the unstable *e*: *douch* (douche) ; 2) *ôtel* (hôtel)
Spanish: N.A.; 2) *emos* (hemos), *rmanos* (hermanos) : normally, "r" is pronounced "erre" (+ phonetic variation); *n* (en), *l* (el) (+ phonetic variation)
Italian: N.A.; not encountered in corpus, all mute "h" forms apparent: *ha, ho, hai, hanno*

2. 1) consonant contractions/clippings & abbreviations; 2) double consonants; 3) semantic abbreviations/initialisms (abbreviations reduced to initials)

French: 1) *dc* (donc), *pr* (pour), *ds* (dans); 2) *ele* (elle), *poura* (pourra); 3) *t* (te/tu) *p* (peux/pas)

Spanish: 1) mñn (mañana), pq (porque), *hcr* (hacer), *td* (todo); 2) not encountered in corpus; 3) q (que/qué), t (te/tu)

Italian: 1) *dmn* (domani), *tnt* (tanto), *scs* (scusa), *ttt* (tutto); 2) not encountered in corpus; 3) *c* (ce/ci), *s* (si/stare)

3. agglutinations

French: *jattends* (j'attends), *jy* (j'y), *na*i (n'ai), *ten* (t'en), *lascenseur* (l'ascenseur); *jmange* (je mange)
Spanish: N.A. for apostrope; *dtus* (de tus), *poray* (por ahi — ex. with phon. Variation), *yokese* (yo qué sé), *noseke* (no sé qué) — ex. with phon. variation & diacritic réduction.

Italian: all apostrophes maintained in corpus; phenomenon could exist: *lho* (l'ho) — private communication

Table 3 (part 2): typology of simple eSMS phenomena across three languages: French, Spanish, Italian

III. suppression/absence or rarefying

1. graphical

1. typography & punctuation

French, Spanish, Italian: in general, typography & punctuation are diminished. In Spanish, initial punctuation (¡, ¿) is systematically suppressed.

2. diacritic signs

French: *ca* (ça), *voila* (voilà)
Spanish: suppression can lead to ambiguity between interrogative & indirect exclamations, so context is required: *quien/quién* (who), *donde/dónde* (where), *cuando/cuándo* (when)

Italian: accents are less frequent. Sometimes suppressed: *ne* (né), but not systematically; "*è"* is quite frequent in corpus

IV. augmentations & additions

1. graphical

1. repetition of characters and/or punctuation

French: *suuuupppeeerrr !!!!!!* (super !)

Spanish: h*olaaaaaaaaa* (hola)

Italian: *arrivooooooo* (arrivo)

2. semiological representations (Smileys/emoticons)

French, Spanish, Italian: :-), :-D, etc.

3. character increase

French: *oki* (ok), *les zamours* (les amours)
Spanish: not encountered in corpus (*okis* for *ok* used in Andalucian, private communication)
Italian: not encountered in isolation; only combined with agglutination — *sennò* (se no)

4. onomatopoeia

French: *mouarf, arfff, bof*

Spanish: *zzz*

Italian: *mhmh, ahahahah*

Table 3 (part 3): typology of simple eSMS phenomena across three languages: French, Spanish, Italian

3. Conclusion

In the present article, we have presented a typology for describing txtng (or eSMS for "écriture de type SMS", or "écrit SMS", SMS-writing). It was initiated for a study conducted on the French language (Panckhurst, 2009), following on from previous research (Anis, 2004, Fairon *et al.* 2006b, Liénard, 2007, Véronis et Guimier de Neef, 2006). Student SMS data was then compiled for Spanish and Italian (Boudrique *et al,* 2008; Stabile & Tortorella, 2008), allowing plurilingual verification in order to reveal any variations, and thus modify typology structure, within the context of intercultural pragmatics. The typology presented here details "simple phenomena" before introducing "complex phenomena" (*cf.* student work, Fréjaville & Doston, 2009). It was first necessary to understand how eSMS forms are devised, for each of the three languages, by carefully splitting up each word into (relatively) isolated linguistic phenomena. However, most words used in text messages combine several features of the above categories (see tables 2 & 3). Although categories vary slightly from one language to another, many features are multilingual & multicultural. However, other languages incorporating other alphabets also need to be researched. For instance, in Arabic, the first letter of the name *Ayesha*/*Aïcha* "عايشة" is "ع". As this is similar to a mirrored letter "3", this number is often used as graphical replacement in text messages. Many more examples need to be investigated further.

Determining typology is a crucial step, before one can branch out and explore "dialogical interactions" or asynchronous SMS "conversations" in depth. In order to do so, a crucial factor is of course collecting sufficient data in interactional situations, which is itself a difficulty in this domain. More student txtng corpora has recently been compiled and analysed (Catapano, 2009, Fontaine, 2009) in order to determine how sSMS interaction takes place (usage of anaphora, ambiguity, discursive variation, terms of address, verb tense usage, contextual situations, etc.). A further step will lead to more precise intercultural & interactional pragmatics issues with analysis of multilingual & multicultural corpora.

4. References

Anis, J. (1999), Internet, communication et langue française, Paris, Hermès.

Anis J. (2001), *Parlez-vous texto ? Guide des nouveaux langages du réseau*, Paris : Le Cherche Midi.

Anis J., de Fornel M., Franckel B. (organisers, 2004) « La communication électronique : Approches linguistiques et anthropologiques », International colloquium, EHESS, Paris, 5-6 February.

Boudrique D., Catapano H., Pollet S. (2008), « SMS espagnols & typologie des SMS français », Master's Assignment, supervisor: R. Panckhurst, Université Paul-Valéry Montpellier 3.

Catapano H. (2009), « Conversations et SMS », Master's thesis, supervisor: R. Panckhurst, Département Sciences du Langage, Université Paul-Valéry Montpellier 3.

Crystal D. (2001), *Language and the Internet,* Cambridge, Cambridge University Press.

Crystal D. (2008*), txting the gr8 db8*, Oxford, Oxford University Press.

Fairon C., Klein J.-R., Paumier S., (2006a), *SMS pour la science. Corpus de 30.000 SMS et logiciel de consultation*, Presses universitaires de Louvain, Louvain-la-Neuve, Manuel+CD-Rom, http://www.smspourlascience.be/

Fairon C., Klein J.-R., Paumier S., (2006b), *Le langage SMS. Étude d'un corpus informatisé à partir de l'enquête « Faites don de vos SMS à la science »,* Presses universitaire de Louvain, Louvain-la-Neuve, http://www.smspourlascience.be/

Fairon C., Klein J.-R., Paumier S., (2007), « Un corpus transcrit de 30 000 SMS français », in Gerbault, J. (éd.) *La langue du cyberespace : de la diversité aux normes*, Paris : l'Harmattan, p. 173-182, http://www.smspourlascience.be/

Fayada Mélanie (2007), *Le langage SMS des jeunes. Approche lexicale et morpho-syntaxique*, Master's thesis, supervisor: R. Panckhurst, Gestion des connaissances, apprentissages, formation ouverte et à distance, Département Sciences du Langage, Université Paul-Valéry Montpellier 3.

Fontaine D. (2009), « Conversations asynchrones en communication médiée : comparaison entre courriels, forums et SMS », Master's thesis, supervisor: R. Panckhurst, Département Sciences du Langage, Université Paul-Valéry Montpellier 3.

Fréjaville D., Doston S. (2009), « Étude des réductions dans un corpus personnel de SMS », Master's Assignment, supervisor: R. Panckhurst, Département Sciences du Langage, Université Paul-Valéry Montpellier 3.

Herring, S. C. (ed), (1996), *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives,* Amsterdam/Philadelphia: John Benjamins.

Liénard F. (2007), « Analyse linguistique et sociopragmatique de l'écriture électronique. Le cas du SMS tchaté », in Gerbault, J. (éd.) *La langue du cyberespace : de la diversité aux normes,* Paris : l'Harmattan, p. 265-278.

Panckhurst R. (1998), « Marques typiques et ratages en communication médiée par ordinateur », Proceedings, CIDE 98, INPT, Rabat, 15-17/04/98, Paris, Europia Productions, 31-43.

Panckhurst, R. (1999), « Analyse linguistique assistée par ordinateur du courriel », in *Internet, communication et langue française*, Anis, J. (ed.), Hermès, 55-70.

Panckhurst R. (2006a), « Le discours électronique médié: bilan et perspectives », in A. Piolat (ed.). *Lire, écrire, communiquer et apprendre avec Internet.* Marseille: Éditions Solal.

Panckhurst R. (2006b), "Mediated electronic discourse and computational linguistic analysis : improving learning through choice of effective communication methods." Proceedings, ascilite conference, p. 633-637, Sydney, Australia, 3-6 Décember.

Panckhurst R. (2007), « Discours électronique médié : quelle évolution depuis une décennie ? », in Gerbault, J. (éd.) *La langue du cyberespace : de la diversité aux normes,* Paris : l'Harmattan, p. 121-136.

Panckhurst R. (2009), « Short Message Service (SMS) : typologie et problématiques futures », in Arnavielle T. (coord.), *Polyphonies*, pour Michelle Lanvin, Université Paul-Valéry Montpellier 3, p. 34-52.

Panckhurst R., Bouguerra T. (2003), "Communicational and methodological/linguistic strategies using electronic mail in a French University", Proceedings, 8th International Symposium on Social Communication, Santiago de Cuba, 20-24 january, 548-554.

Piolat A. (2006, ed.). *Lire, écrire, communiquer et apprendre avec Internet*. Marseille : Éditions Solal.

Stabile T., Tortorella E. (2008), « Le langage de SMS : entre italien et français », Master's Assignment, supervisor: R. Panckhurst, Université Paul-Valéry Montpellier 3.

Véronis, J., & Guimier de Neef, E. (2006), « Le traitement des nouvelles formes de communication écrite », in Sabah, G. (ed*.), Compréhension automatique des langues et interaction*, Paris : Hermès Science.